DOCUMENT RESUME

ED 319 657 SO 020 860

AUTHOR Carifio, James; Lanza, Marilyn
TITLE On the Need to Use Control Vignettes in Social Science Research.
PUB DATE Apr 89
NOTE 32p.; Paper presented at the Annual Meeting of the New England Research Organization (Portsmouth, NH, April 1989).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Control Groups; Females; Males; Nurses; *Research Design; Research Methodology; *Social Science Research
IDENTIFIERS *Vignettes

ABSTRACT

A review of the literature revealed that control (or placebo) vignettes are not used in numerous studies in all areas that employ vignette-based methodology, which is a serious methodological and theoretical flaw. The nature of this flaw and its methodological and theoretical importance is outlined in this paper, as well as procedures and a model for developing control vignettes. Two studies on the degree to which nurse professionals are perceived to be at fault when assaulted by patients are reported which employ control vignettes. In the first study, the use of control vignettes produced better and more objective validation of the assault vignettes than previously available in the literature. In the second study, female and male nurses were blamed as much for the incident that occurred with the patient in the control vignette as they were in the severe assault vignette, but they were blamed more in the mild assault vignette than they were in either the control or severe assault vignette. Female nurses, however, were blamed more than male nurses for the incident in the control and severe assault vignette, but not in the mild assault vignette. These complex findings are shown to fit the non-linear elliptic umbilic model of catastrophe theory, and not the linear models and interpretations of these variables currently in the literature. These significant and important results would not have been found if control vignettes had not been used. (Author)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED319657

# ON THE NEED TO USE CONTROLL VIGNETTES IN SOCIAL SCIENCE RESEARCH

James Carifio, University of Lowell
Marilyn Lanza, Bedford Veterans Hospital

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JAMES CARIFIO

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SO 020 840



BEST COPY AVAILABLE

# ON THE NEED TO USE CONTROL VIGNETTES IN SOCIAL SCIENCE RESEARCH

James Carifio, University of Lowell
Marilyn Lanza, Beford Veterans Hospital

## Abstract

A review of the literature revealed that control (or placebo) vignettes are not used in numerous studies in all areas that employ vignette-based methodology, which is a serious methodological and theoretical flaw. The nature of this flaw and its methodological and theoretical importance is outlined in this paper, as well as procedures and a model for developing control vignettes.

Two studies on the degree to which nurse professionals are perceived to be at fault (blamed) when assaulted by patients are reported which employ control vignettes (verbal aggression only) and mild and severe assault vignettes. In the first study, the use of control vignettes produced better and more objective validation of the assault vignettes than previously available in the literature.

In the second study, female and male nurses were blamed as much for the incident that occurred with the patient in the control vignettes as they were in the severe assault vignette, but they were blamed more in the mild assault vignette than they were in either the control or severe assault vignette. Female nurses, however, were blamed more than male nurses for the incident in the control and severe assault vignette, but not in the mild assault vignette. These complex findings are shown to fit the non-linear elliptic umbilic model of catastrophe theory (Zeeman, 1976), and not the linear models and interpretations of these variables currently in the literature. These significant and important results would not have been found if control vignettes had not been used, which vividly demonstrates the methodological and theoretical importance of control vignettes to research in all social science areas, whether quantitative or qualitative in nature. These findings also call into question the findings of a wide range of vignette-based research in such area as cognitive development (Piaget, 1972), moral development (Kohlberg, 1973) and intelligence (Sternberg, 1984) and suggest that these phenomena are also most probably non-linear in nature. The implication of these studies and the need for replication are discussed. The findings of these studies are pertinent to juries, teacher-student and parent-child relationships, and to the evaluation of professionals.

---

Paper presented at the annual meeting of the New England Research Organization, Portsmouth N.H., April, 1989.

The use of vignettes as a method of eliciting subject responses is becoming more widely used in all areas of social science research and education (Shelley, 1984). Usually, a vignette is a carefully constructed description of a particular situation or problem to which subjects respond in a variety of ways. Typically, conversations, behaviors, surroundings, and other such details are described and a variety of pertinent information is given, and various features or details of the vignette are systematically varied to create different treatment conditions. In some cases, vignettes may be portrayed in a "live" fashion using audio and/or visual representation or even computer animation. Whether in written, audio, audio-visual or computer animated form, however, the purpose of a vignette is to convey a controlled set of circumstances and conditions about which people make judgements and answer questions. A vignette, therefore, is a method for simulating reality; namely, complex events and/or problems. As such, it is an artifical method of providing a controlled stimulus to a wide variety of subjects.

## Strengths and Weaknesses of Vignettes

One advantage of a vignette is that the situation under investigation can be standardized; namely, everyone responds to the same stimulus. As a result, much more control of

extraneous variables is achieved through vignettes as compared to other more open-ended and much used methodologies such as free or structured recall, logs, interviews, observational records, or histories. The vignette methodology, therefore, is a useful technique for controlling otherwise nonrandomized and extraneous behaviors towards which simulated behaviors, intentions to behave, or attitudes are expressed (Shelly, 1984).

Researchers can manipulate variables of interest through vignettes; for example, by varying the degree of intensity of a particular situation or series of events. Subjects can also be assigned randomly to vignettes or sets of vignettes (i.e., treatment groups), and they can be assigned systematically to groups on the basis of a given variable such as sex. A vignette , therefore, is a highly practical and efficient method for collecting a wide variety of data and data of a certain kind. For example, one does not have to wait for the event to occur in an actual situation such as an assault or similar such events to gather data on a post-hoc and uncontrolled basis. Differential recall and other such factors are not problems with vignettes because everyone is responding to the same stimulus with the same time lapse for a response and the particular "event" can even be repeated and repeated over time for longitudinal study. Vignettes, however, do have their weaknesses as a methodology.

Subjects' responses to vignettes cannot be assumed to be identical to their responses to the actual event. One of the major problems in using a vignette is that the stimulus is artifical. Subjects are asked to imagine how they would respond to a particular situation. The actual situation (say an assault) might yield quite different responses. Because of this particular limitation, generalization of findings to the actual situation must be tentative initially, and the external validity of the vignette must be established prior to its use (Flaskerud, 1979). There are, however, internal validity issues as well with vignettes which researchers who use the vignette methodology have ignored to date.

Control Vignettes

To achieve a high degree of internal validity, a researcher usually tries to use an experimental design which includes experimental manipulation of the independent variable, a control group, and random assignment of subjects. A Medline survey of the literature we conducted, however, found that comparison (different amounts of treatment) rather than strict no treatment control groups are used in vignette-based reasearch. Other less formal literature reviews in several other areas also yielded the same results. Not using strict control groups or their equivalent in vignette-based research, therefore, is a major methodological weakness in this body of research, as it leaves a variety of factors uncontrolled and moot in the research, such as response set,

response bias, social desirability, induced response set, halos, levels of baseline responses and other such factors.

The use of a control group or its logical equivalent (an important criterion for a sound experimental design) is somewhat difficult to achieve with vignette methodology, but it can be achieved through construction and use of control vignettes. The control vignette is similar to the other vignettes to which subjects respond with the exception that the independent variable being investigated is absent from the control vignette rather than present to some differing degree as in comparison vignettes. A control vignette, therefore, is the logical equivalent to a placebo. Consequently, the use of control vignettes, like placebos, not only greatly enhances the internal and thus external validity of the investigation, but they also allow one to study the factors that influence the misclassifications and misinterpretations of the control vignettes and thus indirectly the target vignettes as well.

## Constructing Control Vignettes

Constructing control vignettes is not as easy or straight forward a task as it might seem to be. How do you ask a subject to respond to an event that has not occurred (or a non-problem or non-dilemma)? It is indeed a challenging problem, and in a recent study, we had to face this challenge. In our study, we were investigating the physical assault of nurses by patients through use of 4 vignettes that

had been previously constructed and validated (Lanza, 1987). These 4 vignettes varied the sex of the nurse (female or male) and the severity (mild or severe) of the physical assault by the patient, who was a male in all 4 vignettes.

After reading one of these vignettes, subjects would answer 13 questions using a five point rating scale which assessed the degree to which the professional nurse was responsible (to blame) for the assault that occurred. Five of the 13 items required reverse scoring to derive a total blame score. Our task, therefore, was to create a control vignette where no assault occurred and questions could be asked about this situation that were the same as the other "target" vignettes. To do this, we found that we had to more abstractly and rigorously define our dependent variable (assault) and the surrounding conditions in which the assault occurred or did not occur. Severity of assault was defined as the degree of force used by the patient on the nurse and the degree of physical injury to the nurse. Severity of assault was simply classified as severe or mild in the present instance.

To have a control vignette, the subjects would have to have something to which they could respond that was similar to the independent variable being investigated, but the independent variable could not be present in the control vignette. After several unsuccessful efforts, we realized that physical assault was a subcategory of a more general variable, aggressive behavior, which could manifest itself in

verbal and non-verbal forms. We also began to think about each vignette situation as an "event" rather than as an "assault."

The control vignette we constructed described a situation and events similar to the target vignettes, but physical assault did not occur in the control vignette. The control vignette included the same conversations and behaviors as the other vignettes, but the outcome was verbal anger only on the patient's part. There was no physical assault. The original questions following each version of these vignettes had to be changed so that they would be the same and work appropriately across all of the versions of the vignette. For example, the word "incident" rather than "assault" had to be used to refer to events. Therefore, by developing a much more abstract and generalized "blueprint" or construct of our dependent variable, associated variables, and events, we were able to construct a control vignette without a great deal of difficulty and could in fact construct several such control vignettes. Consequently, since we abstracted a model, the vignettes we actually used in the present study were a sample of the possible vignettes that could have been devised to investigate aggressive behaviors and reactions, and the model was a general vignette "blue-print" that could be used to make an audio, audio-visual, or computer animated presentation of the vignettes for comparison to written presentations.

Description of the Vignettes Used

Basically, the vignettes used in the present study described a somewhat casual conversation about a patient's weekend pass between a nurse (female or male) and a patient (male) who occasionally hit people. The patient indvertently tells the nurse that he is going home on a two day weekend pass. It is the nurse's understanding, however, that a one day pass has been approved for the patient by his treatment team, and when this fact is related to the patient, the patient becomes hostile and verbally aggressive towards the nurse. The nurse then tries to calm the patient down by suggesting that they discuss the matter.

The three different treatment conditions used in the present study varied in terms of the ending used to this vignette; namely, in terms of what happened from this point forward in the vignette.

In the Control (verbal aggression only) vignette the patient tells the nurse aggressively what the nurse "can do" and where the nurse "can go," and then abruptly turns and walks away, ending the vignette.

In the Mild Assault vignette, the patient aggressively tells the nurse what the nurse "can do" and where the nurse "can go," and then grabs the nurse very hard by the wrist and will not let go. The nurse needs help to get free from the patient and the nurse's wrist has a reddened mark on it after the incident is all over, which is the end of the vignette.

In the Severe Assault condition, the patient aggressively tells the nurse what the nurse "can do" and where the nurse "can go," and then grabs the nurse very hard by the wrist and starts punching the nurse on the arm. The nurse loses her or his balance, falls, and hits her or his head against the wall. When help arrives, the nurse is bleeding from several head cuts and has a severely sprained arm and wrist, which is the end of the vignette.

As previously stated, after reading one of the above vignettes, a subject would answer 13 questions, using a five point rating scale, which assessed the degree to which the professional nurse was responsible (to blame) for the incident that occurred. These 13 items, which were exactly the same for all vignettes, were found to have two underlying, but correlated, factors (using an eigen cut-off of 1.0) that accounted for 72% of the variance observed in the sample (N=58). These two factors were Personal Blame and Blame as Perceived by Fellow Workers. As the two factors were correlated (r=+.52), a Total Blame score, which was a simple summation of all 13 items, was used in most analyses conducted. The Cronbach internal consistency coefficient for these 13 items was r=+.91 (N=64), and the one week test-retest reliability coefficent was r+.86 (N=55). These reliability coefficients were the same across all forms of the vignettes used and were not different from those observed by Lanza (1987) for the original 4 vignettes.

All 13 items significantly predicted total blame score

with both the median and mean item total correlation being r=+.70 (N=64). Item total correlations ranged from .36 to .89 with 10 of the 13 item total correlations being above .70. Total Blame scores, moreover, were also found to correlate with age at r=+.31 (N=58,p.<.01), job experience at r=+.20 (N=58,p.<.10) and with score from the Rubin and Peplau (1973) Just World Scale at r=+.54 (N=55,p.<.01). The level of these concurrent validity coefficients were about as predicted from theory and the literature. The Total Blame scale and scores used in the present study, therefore, were quite good psychometrically.

Methodology and Results

To establish the external, objective, or content validity of the 2 control vignette and the other 4 vignettes used in the present study, we used a panel of 12 professional experts and a modified version of the Yudofsky (Yudofsky et al, 1986) Overt Aggression Scale for the experts to rate the degree to which specific vignette characteristics were present or absent in the 6 vignettes. We presented all vignettes to each rater in a different random order. The raters completed the modified Yudofsky scale for a vignette before responding to the next vignette in her or his package. The reason we used the Yudosky Scale is that it is used clinically and administratively in the assessment of aggressive behavior and is one of the few scales in existence and use, and we wished to further assess its usefulness and

relate our data to the literature in this area.

The Yudofsky Overt Aggression Scale (Yudofsky, 1986) has four major categories of aggressive behavior (verbal aggression, and aggression against self, objects and others). Each category has 4 to 5 general, well-described, observable behaviors that are exemplars of the category, which are arranged in order of increasing severity of exemplification. Yudofsky reports good test-retest reliability (r=+.87) and inter-rater (r=+.75) and validity for the scale in general. Yudofsky, however, found that users of his scale varied considerably on the specific items and categories marked when rating the same person or situation, but not in terms of global summations raters were required to write about the person or situation rated (see Yudofsky for further details).

The modifications made to the Yudofsky scale in the present study were (1) only 2 of Yudofsky"s 4 overt aggression categories were used (verbal aggression and physical aggression towards others), and (2) a global summative judgement item was added to allow the rater to classify the vignette overall as being a vignette that portrayed verbal aggression or mild or severe assault. The global item was followed by 6 items that asked the expert raters to judge the plausibility of the vignette in terms of various characteristics and features concerning the location of the incident and conversations and behaviors of the nurse and the patient in the vignette.

The panel of expert raters were 3 physicians, 2

psychologists, and 1 social worker, all of whom were male, and 6 nurses, 5 of whom were female and 1 of whom was male. The average age of the raters was 43 years, and the average number of years of hospital work experience was 16 years. Nine of the 12 raters had experienced mild assault on the job, and 2 had experienced severe assault.

In terms of the global vignette type classification measure added to the Yudofsky scale, 100% of the 12 raters accurately identified both of the control vignetttes as verbal agression only and both of the mild assault vignettes as mild assault. Only 9 of 12 (67%) of the raters, however, correctly identified the severe assault vignettes as severe. The three raters who misclassified the severe assault vignettes were very "assault experienced" raters and made up their own category of "moderately severe" for these vignettes which they wrote on the rating scale and checked. This finding is not only somewhat startling, given the vignettes described above, but says more, we believe, about the level of severity of assaults that occur in hospitals and the degree to which professionals become inured to them, than the actual characteristics of the severe assault vignettes. That the 3 misclassifiers were male raters rathan than female raters was significant, we believe, as will be seen from the data in the main experiment reported below.

The inter-rater reliability coefficient for the global classification of the vignettes was r=+.94, and, in general, all 12 of the raters found all of the vignettes to be

plausible on the 6 factors rated. At the global level, therefore, the evidence for the objective validity of the vignettes constructed was very good.

In terms of the 9 items and 2 categories of the modified Yudofsky scale itself, we devised two procedures for devising a total scale and subscale scores for each rater on each vignette. As the 9 items of the modified Yudofsky scale were arranged in order of increasing aggressive behaviors, we successively added one additional scoring point to each item to give the item a weight that reflected the increasing severity in aggressive behavior that it represented as one progressed up the scale. We also used this weighted scoring procedure to derive total subscale scores both within and between the two categories of the scale so that each category could be examined both individually and comparatively. The weighted scoring system that we devised to derive a total severity of aggression present in the vignette score and a degree of physical assault present in the vignette subscale score (the second category of the scale) is a very conservative weighted scoring system that is conceptually similar to the weighted scoring procedure Yudofsky reports using in his validity studies of the scale. The reason that we devised the weighted scoring procedures reported here, however, is that Yudofsky did not report the scoring weights he used for his scale in his studies. Lastly, it should be noted that one male rater who checked 8 of the 9 items on the scale for all vignettes was dropped from the analyses

conducted, as this rater was so extremely deviant and illogical, as compared to the other 11 raters who were relatively consistent. This rater, moreover, was one of the 3 raters who made up their own "moderately severe" global scoring category for the severe assault vignette. In terms of the 11 raters retained for analyses, their internal consistency in ratings across the 9 items was r=+.56. This internal consistency coefficient, which is reasonable for 9 items of this type and mixture, was established by conducting a Severity of Incident with the Patient by Sex of the Nurse Practitioner (3x2) Hoyt repeated measures analysis of variance as described in Kerlinger (1986).

Table 1 presents the results of a Severity of the Incident with the Patient by Sex of the Nurse Professional in the Vignette (3x2) repeated measures ANOVA for the Total Severity of Aggression in the Vignette weighted score derived for 11 of the 12 raters in this vignette validation study. As can be seen from Table 1, a highly significant main effect difference in total severity of aggression in the vignette was observed between the control, mild and severe vignettes, with roughly ratio level differences between the means of each vignette type, which should be present if each vignette type represents a truly different level of severity in aggressive behaviors. No significant differences were found relative to the sex of the nurse professional in the vignette or significant interaction. Raters found no differences in the total severity of aggresion present in a given vignette

Table 1: SEVERITY OF INCIDENT WITH PATIENT BY SEX OF THE NURSE PROFESSIONAL INVOLVED (3X2) REPEATED MEASURES ANOVA ON TOTAL SEVERITY OF AGGRESSION IN THE VIGNETTE WEIGHTED SCORES (N=11).

Severity of Incident in Vignette

| Sex of Nurse Prof. in Vignette | (Verbal Abuse) Control N | Mean | St.D. | Mild Assault N | Mean | St.D. | Severe Assault N | Mean | St.D. |
|---|---|---|---|---|---|---|---|---|---|
| Female Nurse | 11 | 3.5 | 1.9 | 11 | 8.5 | 3.7 | 11 | 14.2 | 6.0 |
| Male Nurse | 11 | 3.3 | 1.8 | 11 | 9.0 | 4.2 | 11 | 15.3 | 6.7 |
| Total | 11 | 3.4 | 1.8 | 11 | 8.7 | 3.9 | 11 | 14.7 | 6.4 |

| Source | df | Mean Sq. | F | p |
|---|---|---|---|---|
| Subjects | 10 | 40.91 | | |
| Sev. of Incident(SI) | 2 | 710.97 | 23.17 | <.0001* |
| Error | 20 | 30.68 | | |
| Sex of Nurse (Sex) | 1 | 3.88 | 1.18 | >.05 |
| Error | 10 | 3.28 | | |
| SI x Sex | 2 | 2.24 | 0.35 | >.05 |
| Error | 20 | 6.39 | | |
| Total | 65 | | | |

Table 2: SEVERITY OF INCIDENT WITH PATIENT BY SEX OF THE NURSE PROFESSIONAL INVOLVED (3X2) REPEATED MEASURES ANOVA ON DEGREE OF PHYSICAL ASSAULT PRESENT IN THE VIGNETTE WEIGHTED SUBSCALE SCORES (N=11).

Severity of Incident in Vignette

| Sex of Nurse Prof. in Vignette | (Verbal Abuse) Control N | Mean | St.D. | Mild Assault N | Mean | St.D. | Severe Assault N | Mean | St.D. |
|---|---|---|---|---|---|---|---|---|---|
| Female Nurse | 11 | 0.0 | 0.0 | 11 | 1.4 | 0.9 | 11 | 3.3 | 1.6 |
| Male Nurse | 11 | 0.0 | 0.0 | 11 | 1.4 | 0.9 | 11 | 3.1 | 1.4 |
| Total | 11 | 0.0 | 0.0 | 11 | 1.4 | 0.9 | 11 | 3.2 | 1.5 |

| Source | df | Mean Sq. | F | p |
|---|---|---|---|---|
| Subjects | 10 | 2.05 | | |
| Sev. of Incident(SI) | 2 | 56.06 | 28.12 | <.0001* |
| Error | 20 | 1.99 | | |
| Sex of Nurse (Sex) | 1 | 0.61 | 1.00 | >.05 |
| Error | 10 | 0.61 | | |
| SI x Sex | 2 | 0.61 | 1.00 | >.05 |
| Error | 20 | 0.61 | | |
| Total | 65 | | | |

in terms of the nurse practitioner in the vignette being male or female.

A clearer understanding of these results presented above may be obtained by examining the degree of physical assault present in the vignette weighted subscale score that was devised. Table 2 presents the results of a Severity of the Incident with the Patient by Sex of the Nurse Professional in the Vignette (3x2) repeated measures ANOVA for the Degree of Physical Assault Present in the Vignette weighted subscale score for 11 of the 12 raters in this vignette validation study. As can be seen from Table 1, a highly significant main effect difference in total severity of aggression in the vignette was observed between the control, mild and severe vignettes, with greater than ratio level differences between the means of each vingette type, which should be present if each vignette type represents a truly different level of severity in aggressive behaviors. No significant differences were found relative to the sex of the nurse professional in the vignette or significant interaction. Further, as can be seen from Table 2, no phsyical assault was found in any of the vignettes by any of the 11 raters using the Yudosky scale.

From the results presented above, it would seem reasonable to say that the item-level evidence for the objective and content validity of the 6 vignettes constructed is more than relatively good. Using the modified Yudofsky scale, subjects could objectively and reliably discriminate

out levels of aggressive behaviors and assault and non-assault behaviors in the vignettes, with ratio level differences between the means of each vignette type. The ability of subjects to make these kinds of discriminations and to make them reliably has been a major and outstanding unanswered question in the literature in this area. However, this question would not have been answered as well or as convincingly as it is in this study if we had not used control vignetttes in this study. Furthermore, that this particular question is well answered is an important point in terms of the second study we conducted.

In a second experiment, 64 adult subjects were randomly assigned one vignette to read and answer questions on two occasions a week apart. The second administration of the same vignette to each of these subjects was to establish the test-retest reliability coefficients of the control and rewritten vignettes, which were reported above. The results of the first administration of the vignettes were used in the analyses that will be reported in this paper and the results of the second administration were used to cross validate these results. Subjects were asked to respond to only one of the six vignettes, rather than 3 of one sex or the other, or all 6, to simplify the experimental design and analysis and to ensure that responses to the control vignettes were not in any way affected by reading other vignettes beforehand. Once the performance of the control vignettes were clearly ascertained under neutral conditions, therefore, the ground

work for more complex experiments and experimental designs would be established.

The subjects in the second experiment were nurses working in a veterans administration hospital. All of the nurses in this hospital were randomly sent one the the six vignettes to read and respond to using the 13 item blame scale. These nurses were then asked to write an explanation of their responses to the vignettes, and then to fill out a background questionnaire and to respond to the Just World Scale. Of the 251 nurses initially sent one vignette and the associated material just described, 64 (25.5%) returned the package sent. Of the returners, 58 were female and 6 were male subjects. As there were no male subjects for some of the cells in the design, and the general responss patterns of the 6 male subjects seemed to be different from the female subjects, the male subjects were dropped from the analyses reported below.

The response rates for each of the 6 vignettes were not the same. Only 16 subjects returned the control vignettes, while 17 subjects returned the mild assault vignette and 25 subjects returned the severe assault vignettes. In order to assess possible biases in these differential response rates, a Severity of Assault by Sex of the Nurse Professional (3x2) MANOVA was run on all the background variables for the 58 female subjects that were to be used in analyses. No significant differences were found between subjects on any of the 7 variables examined. These analyses, therefore, tend to

indicate that the differential response rates most probably do not indicate bias and that the sample and the assignment of vignettes process was mostly probably random. Further evidence will be presented below to support this view. The mean age of the 58 female subjects used in analyses was 45.7 years with a standard deviation of 9.9 years, and the mean number of years of experience was 18.5 with a standard deviation of 9.9 years. These 58 subjects also varied significantly in terms of job group level in the hospital and level of education. Lastly, it should be noted that 58 subjects in a 2x3 ANOVA resulted in some small (<10) cell sizes, and that this fact needs to be taken into account in interpreting results.

Table 3 presents the results of a Severity of the Incident with the Patient by Sex of the Nurse Professional in the Vignette (3x2) ANOVA for the Total Degree to which the Nurse Professional in the Vignette is Blamed score for the 58 female subjects in the present study. As can be seen from Table 3, female and male nurses were blamed as much for the incident that occurred with the patient in the control vignettes as they were in the severe assault vignette, but they were blamed more in the mild assault vignette than they were in either the control or severe assault vignette. Female nurses, however, were blamed more than male nurses for the incident in the control ($p<.13$) and severe assault ($p<.05$) vignette, but not in the mild assault vignette. These results are, without doubt, truly remarkable, and

Table 3: SEVERITY OF INCIDENT WITH PATIENT BY SEX OF THE NURSE PROFESSIONAL INVOLVED (3X2) ANOVA ON TOTAL BLAME OF THE NURSE FOR THE INCIDENT SCORES (N=58).

Severity of Incident in Vignette

| Sex of Nurse Prof. in Vignette | (Verbal Abuse) Control N | Mean | St.D. | Mild Assault N | Mean | St.D. | Severe Assault N | Mean | St.D. |
|---|---|---|---|---|---|---|---|---|---|
| Female Nurse | 9 | 35.2 | 9.4 | 7 | 38.4 | 12.6 | 13 | 34.1 | 7.5 |
| Male Nurse | 8 | 29.6 | 11 | 9 | 39.8 | 9.7 | 12 | 29.6 | 9.4 |
| Total | 17 | 32.6 | 10.3 | 16 | 39.3 | 10.7 | 25 | 32.0 | 9.4 |

| Source | df | Mean Sq. | F | p |
|---|---|---|---|---|
| Sev. of Incident(SI) | 2 | 288.76 | 3.04 | <.056* |
| Sex of Nurse (Sex) | 1 | 102.22 | 1.08 | >.05 |
| SI x Sex | 2 | 80.11 | 0.84 | >.05 |
| Error | 52 | 94.83 | | |
| Total | 57 | | | |

Table 4: SELECTED TREND ANALYSES OF SEVERITY OF INCIDENT WITH PATIENT BY SEX OF THE NURSE PROFESSIONAL INVOLVED (3X2) ANOVA ON TOTAL BLAME OF THE NURSE FOR THE INCIDENT SCORES (N=58).

| Trend Factor | Source | df | Mean Sq. | F | p |
|---|---|---|---|---|---|
| Female Nurse | linear | 1 | 15.79 | 0.10 | >.05 |
| | non-linear | 2 | 15860.00 | 160.33 | <.001* |
| | error | 25 | 98.92 | | |
| Male Nurse | linear | 1 | 9.47 | 0.10 | >.05 |
| | non-linear | 2 | 17633.00 | 192.60 | <.001* |
| | error | 25 | 91.52 | | |
| Both Combined | linear | 1 | 1.17 | 0.02 | >.05 |
| | non-linear | 2 | 34190.00 | 517.50 | <.001* |
| | error | 55 | 66.07 | | |

contrary to the conventional theories, views, and data reported in this area.

The results present in Table 3 are not insignificant, or marginally significant, as indicated by the F-ratios in the table, but rather they are highly significant and are only observable because control vignettes were used. The data present in Table 3, contrary to the F-ratios in the table, are highly significant, predictable and "regular", as can be seen from the trend analyses results presented in Table 4 where the F-ratios exceed 100. The reason that these results are so highly significant is that the pattern of the data in Table 1 is what is called "linear chaos" (Gleick, 1988) or a linear models "catastrophe" (Zeeman, 1976).

Essentially, chaos theory (Gleick, 1987) is the view that what might appear to be chaotic, unpredictable, and random from one point of view may have order, structure and predictability from another point of view, and what might appear to be phenomena and variables that may be "well approximated and represented" by linear models are very often not so, and non-linear to such a significant degree that linear representations of these phenomena and variables produce severe distortions and illogicalities in both theory and real world observations. Catastrophe theory is essentially a theory and mathematics of non-linear functions and relationships that is far more specific and precise than chaos theory, which is a general and all emcompassing theory. Non-linearity is a mathematical and theoretical "catastrophe"

for linear models and conceptualizations. A theoretical or observational "catastrophe" occurs when the linear model predicts the "next data point" or results from previous data points and the predicted result in reality is observed to be in exactly the logically and mathematically opposite place.

Seven very distinct and different general "catastrophe" models have been developed (Zeeman, 1976) which range from simple to complex depending on the number of observable factors and behavioral dimensions that are related together in the model. Each of these general models describes how the "regression surface" or "plane" in the linear view is "folded" into contoured shapes or structures that can be described by mathematical equations which make the data orderly, structured, and predictable both theoretically and logically as well as mathematically.

The data present in Table 3 is representative of and conforms to the "front face" of the elliptical umbilic catastrophe theory model presented by Zeeman (1976) for non-linear relationships that have 3 "control" (i.e., Severity levels) and 2 "behavioral" (i.e., sex levels) dimensions, and is most likely the first "real world" experimental instance of data that fits this model in the social sciences. The elliptical umbilic model is an extension of the simple "cusp folded" (one-way) model developed by Zeeman to explain and predict the non-linearity of aggressive behavior in simple situations to a two factor (3x2) situation. Understanding the model and the ways in which it functions, and catastrophe

theory, is not as difficult as it may appear to be, and both the model and its operations may be easily understood in two rather simple and straight forward ways.

The first way to understand the model and catastrophe theory is by reducing the data in the 3x2 table presented to the three 2x2 tables that are inherent in the 3x2 table, and then running anovas on three 2x2 tables. If this were actually done, significant main effects and significant disordinal interactions would be observed for the control to mild and severe to mild incident 2x2 tables, but not the control to severe 2x2 table. In the control to severe 2x2 table, however, significant sex differences would be observed, which is intuitively obvious from the 3x2 table presented. Any one of these 2x2 tables, however, would be a significantly inaccurate and misleading characterization of the phenonenon under investigation and its essential nature and behavior. All of these 2x2 effects and the inherent structure underlying them do not come out in the 3x2 table because they are hidden and masked in the linear assumptions of the model used and all of the analyses conducted with it. However, when the different parts of the whole design are considered, they are clearly and obviously there. The data, variables under investigation and their relationships, therefore, are not chaotic or disordinal. In a word, they are logically complex and non-linear.

The second way to understand catastrophe theory and the data presented in Table 3 is to run trend analyses on the

data for non-linear (quadratic) as well as linear trend. The quadratic trends, if significant, would show that the data are following and conforming to highly predictable non-linear functions. Table 4 presents the results of trend analsyes on the data in table 3. As can be seen from table 4, the linear trends are non-significant, but the non-linear trends are highly significant with F-ratios exceeding 100, which should eliminate any questions one might have about sample size or unequal cell sizes.

The trend analyses clearly show that the data are best characterized as fitting a highly significant and predictable non-linear function. However, it should be noted that only catastrophe theory provides a precise model and explanation of the exact nature, structure and predictability of the non-linear model observed, which is not obtainable from the trend analysis itself. One must have the concepts of "strange attractors," "folded space," and "dimension jumps" due to logical necessity, differential conditions, and the "folds" that relate them together to obtain a cogent and non-empirical model and explanation of the data and the variables under investigation, all of which would not have been possible if control vignettes had not been used, which is one of the central point of this paper. The point is that the phenomenon of blaming and judging is this complex and cannot be "simplified," as such a "simplification" would be an over simplification, which is a point that is true for a wide variety of phenomena, and the central point that catastrophe

theory addresses. We need catastrophe theory, therefore, to understand and model human thinking and human thinking processes, and to make machines and models that are capable of behavior in a similar fashion. In a word, all empirical views of the data presented here and the phenomenon under investigation are incomplete, and to a very great degree completely empty and merely empirical without it.

## Discussion and Conclusions

The data, analyses, and results present in this paper were meant to be illustrative rather than definitive. The purpose of this paper was not to fully analyze or present all of the data or findings in the studies conducted, nor was the purpose to elucidate fully all of the details of chaos or catastrophe theory. Rather, the purpose of this paper was to outline and demonstrate the need for and extraordinary importance of using control vignettes, placebos, and baseline behavior description devices in research, not only with vignettes, but also in a wide variety of other kinds and types of research as well. The use of control vignettes is not only extraordinarily important methodologically, but also important theoretically as well, and in terms of accurate characterization, understanding, and interpretation of the phenomenon under investigation, and findings, as we believe the results present in this paper clearly demonstrate.

The present paper is to some degree theoretical, but it is also highly practical and results orientated at the same

time. The theory presented is not only supported by observed results, but also by innovations in general theoretical perspectives that have been occurring in several different areas over the past 15 years. It is for these reasons that we believe that the findings and views presented in this paper are important, highly significant and correct not only for vignette-based research, but also for a wide variety of other type of research as well.

The present study only answered a limited number of questions about control vignettes and their use, and several further studies were suggested. Using control vignetttes, however, did create a baseline for interpreting the data in our studies and having this baseline produced some rather startling results that would not have been known without the control vignettes, nor would our data have been clearly interpretable without them. In fact without the control vignettes, the data from our studies would have been interpreted quite differently, and this different interpretation would not only have been incorrect, but also highly misleading relative to understanding the phenomenon under investigation, which is essentially a non-linear or "chaotic" phenonemon rather than a "linear" or "regular" phenomenon. This deeper level of understanding would not have come about or emerged if the control vignettes were not used, which is the central point of this paper. Use of control vignettes in our studies, therefore, significantly and qualitatively enhanced our investigation and are needed

to enhance the internal and external validity of vignette investigations. Consequently, unlike the prevailing wisdom, we believe that classical experimental principles and theories are highly pertinent and germane to qualitative research methods, and can easily be applied in most instances to qualitative methodology to improve the methodology and the results obtained.

Relative to vignette-based research, our findings not only call into question the findings of the numerous studies that have been done on patient assault, but they also call into question the findings of a wide range of vignette-based research in such area as cognitive development (Piaget, 1972), moral development (Kohlberg, 1973) and intelligence (Sternberg, 1984), and suggest that these phenomena are also most likely non-linear in nature. The implications of this particular point are not trivial and is one of the reasons why replications of the studies reported here, as well as other types of studies of control vignettes, are needed. Lastly, we believe that the findings of the studies reported here are highly pertinent to juries, teacher-student and parent-child relationships, and to the evaluation of professionals, relative to understanding the process and conditions under which levels of blame are differentially attributed to people depending on the characteristics of the incident and the sex of the person being evaluated. The results of our study were to some degree counter intuitive, but highly supportive of the contentions of many groups that

differential standards are used in the evaluation of their behaviors and performances by others. There is most probably no one who has not observed or experienced first-hand an evaluative or blaming "catastrophes" of the kind depicted in the data presented in this paper, which is a validation of subjective, qualitative experience that would not have emerged and been validated if control vignettes had not been used in the present study.

## References


Cronbach, L. Playing with chaos. Educational Researcher, Sept., 1988, p 46-49.

Gleik, J. Chaos:making sense of a new science. Penguin, 1987.

Flaskerud, J.H. Use of vignettes to elicit responses towards broad concepts. Nursing Research, 28, 1979, 210-211.

Rubin, Z. and Peplau, A. Who believes in a just world? Journal of Social Issues, 31 (3), 1973, 65-89.

Kerlinger, F. Foundations of Behavioral Research. Holt, Rinehart, and Winston, 1986.

Kohlberg, L. The child as moral philosopher. Psychology Today, 2(4), 1973, 25-30.

Lanza, M. The relationship of severity of assault to blame placement of assault. Archives of Psychiatric Nursing, 1(4), 1987, 269-279.

Lerner, M.J. and Miller, D.T. Just world and the attribution process: looking back and ahead. Psychological Bulletin, 85 (5), 1978, 1030-1051.

Piaget, J. Intellectual evolution from adolescence to adulthood. Human Development, 15, 1972, 1-15.

Shelley, S.I. Research methods in nursing and health. Boston, Little Brown Company, 1984.

Sternberg, R. What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. Educational Researcher, Jan, 1984, 5-14

Yudofsky, S.C., Silver, J.M., Jackson W., Endicott, J., and Williams, D. The overt aggression scale for objective rating of verbal and physical aggression. American Journal of Psychiatry, 143(1), 1986, 35-39.

Zeeman, C.E. Catastrophe Theory. Scientific American, 234, 4, April, 1976, 65-83.